

National Center for Education Statistics Disclaimer

The information and opinions published here are the product of the International Indicators of Education Systems project's Network A and do not necessarily represent the policy or views of the Department of Education or the National Center for Education Statistics.



Review of Assessment Activities

In this Issue

It has been a while since the last issue of the newsletter, but we are back—and with a new look! And from now on, we will be distributing the newsletter both electronically and in hard copy, which will be more timely, cost-effective, and hopefully more convenient for you, should you wish to share the newsletter with interested colleagues in your countries.

The feature article of this July 2003 newsletter presents information on the newly released mathematics portion of the TIMSS 1999 Video Study, which examined mathematics and science teaching in seven countries. The article describes both the innovative methodology employed by the study and the results indicating differences in mathematics teaching practices across countries, as well as briefly looks at some of the benefits and challenges in using video in studies of education.

Also included in this issue is a country highlight focusing on the United States. Read about education and assessment in the United States—in particular, about the impacts of *No Child Left Behind* on the well-known National Assessment of Educational Progress (NAEP) and assessment and testing activities at the state level. As usual, the newsletter also provides updates on Networks A, B, and C, and the BPC, as well as a spotlight on the work of the INES Task Force on Teaching and Learning. National assessment practices will be highlighted in the December issue of the newsletter from now on.

We thank all those who contributed to the newsletter, including Marilyn Binkley, from the U.S. National Center for Education Statistics for contributing to the article on assessment in the United States, and Ann-Caroline Nordström of Sweden and Jaap Scheerens and Maria Hendriks of the Netherlands for sharing information on Networks B and C. We appreciate your efforts in keeping us informed of activities from around the INES Project. We hope you enjoy the latest newsletter!

The TIMSS 1999 Video Study of Eighth-Grade Mathematics Teaching

Results were recently released from the mathematics portion of the Trends in International Mathematics and Science Study (TIMSS) 1999 Video Study (results from the science portion will be released at a later date). This study, through in-depth analysis of videotapes of eighth-grade mathematics lessons, examined and compared classroom-teaching practices in seven countries. This article describes both the design and key results of the study, and also, more generally, describes some issues in using video in studies of education.

Benefits and challenges of using video to study teaching

It is safe to assume that across countries and cultures, classroom teaching profoundly affects the ways in which children learn. But while teaching may occur in one form or another almost universally

throughout the world, the specific way in which teaching is practiced certainly varies, at least on some measures, across countries. Studies that examine these cross-country variations in teaching practices can help us better understand education and culture. More specifically, such studies may play an important role in helping educators to effectively identify factors that enhance student-learning opportunities and, by extension, student achievement.

Teaching, however, can be difficult to study and analyze in a way that accurately captures the complexity and nuance of classroom interaction and lesson content – particularly in a cross-cultural context. Traditional studies of classroom teaching have tended to rely primarily on teacher questionnaires, which are relatively easy to administer and provide readily available quantitative data. However, while questionnaires can be useful for gathering information on teachers' backgrounds, percep-

Early examples of the use of video in education studies

One of the earliest comparative studies of education using video technology was conducted by educational anthropologists George and Louise Spindler. These researchers conducted a long-term study (begun in 1977) in two schools in Schoenhausen, Germany and Roseville, Wisconsin, U.S.A., to examine the influence that culture had on the role of the school in preparing children for their rapidly changing social and economic environments. Their study used video to obtain more comprehensive records of the activities and interactions that were going on in the schools and to elicit candid thoughts through interviews with the subjects of the film. The results reinforced the simple but important idea that schools provide a vivid cultural context for students.

The Preschool in Three Cultures project provides another example of a cross-national education study that used video technology. Unlike some other studies, however, video was used not to collect data for analysis, but instead to collect images that could be used as cues for reflection by teachers. This project (Tobin, Wu, and Davidson, 1989) videotaped an entire day in classrooms in one preschool each in China, Japan, and the United States. Subsequently, the researchers played the videotapes for teachers and asked them to explain the thinking behind their words and activities. Some of the country-specific practices observed in the study included that Japanese teachers held back from intervening in children's disputes, Chinese teachers attempted to correct any over-indulgence that only children receive at home, and American teachers emphasized the importance of children expressing their feelings through words.

SOURCE: National Research Council. (2002). *Power of Video Technology in International Comparative Research*. Washington, D.C.: National Academy Press.

tions, plans or working conditions, they are less effective at capturing information on actual teaching practice. This is because they are dependent on teachers' memories and subjective interpretation of the questions, and they also may be limited by the fact that teachers may not be consciously aware of some of their behaviors in the classroom.

By contrast, video studies allow for a more thorough investigation of teaching as it is actually practiced. The use of video facilitates the examination of complex behaviors and circumstances, enables coding of those behaviors and circumstances from multiple perspec-

tives, and preserves classroom activities in such a way that they can be reviewed, slowed down, and stored for future analysis. Furthermore, video creates novel ways to communicate the results of a study, allowing for integration of quantitative and written reports with "real" examples from video cases that can be used for both re-

search and professional development purposes. (The textbox on page 2 describes two early education studies to use video for purposes other than analysis.)

However, while the use of video may address some of the problems associated with traditional methods for studying teaching, it also brings its own set of challenges. For example, to ensure that the individuals charged with coding the observed behaviors have similar views on the classrooms, video studies require that researchers develop explicit standardized camera procedures and train videographers to follow these procedures. Furthermore, due to the high cost of filming, particularly in different countries across the world, there may be significant constraints to the number of cases that can be included in video studies, which places added importance on developing a strong sampling plan. Achieving reliability among coders, especially when they are

examining behaviors or phenomena that are difficult to explicitly define or categorize, is also a significant challenge faced by researchers using video. This challenge is compounded in cross-national studies like TIMSS, when coders often represent many different language groups and cultures. Finally, researchers using video must make conscious efforts to minimize the effect of the presence of the camera on the participants being filmed (observer effects), and to assess, through surveys of participants, what effect the presence of the camera may have had on their results. In other words, there are many organizational, financial, and methodological challenges that must be addressed in the design of video studies.

Design and methodology

The TIMSS 1999 Video Study was conducted under the auspices of the International Associa-

Building from the first TIMSS Video Study

The TIMSS 1999 Video Study was designed to expand upon the first TIMSS Video Study of mathematics teaching practices, which was conducted in 1995 in Germany, Japan, and the United States and which was the first major cross-national study to use video technology specifically to examine and compare teaching practices across countries. From the 1995 study, a key hypothesis emerged: different countries have distinct cultural patterns of teaching mathematics.

Therefore, the 1999 study was designed to explore these patterns in more depth and to address lingering questions that had not been answered due to limits in the scope of the first study. The 1999 study was also able to profit from advances in video methodology that had occurred since the first study. Two of the ways in which the 1999 study differs from the 1995 study are an increase in the number of participating countries and more uniformity in terms of the achievement level of the participating countries

The results from the 1995 study generally indicated that Japan had a distinct pattern of teaching in contrast to the other two countries. Japan was also the highest performing country of the three based on the 1995 TIMSS mathematics assessment. By contrast, the five new countries included in the 1999 video study (Australia, Czech Republic, Hong Kong SAR, the Netherlands, and Switzerland) could each be classified as “high performing” based on the TIMSS 1995 mathematics assessment. While it is important to note that conclusions cannot be drawn from this study about the “superiority” of certain practices over others based on differences in mathematics scores of the countries, the inclusion of many high achieving countries in the 1999 study allows for a general comparison of different ways that teaching is practiced in high achievement settings.

tion for the Evaluation of Educational Achievement (IEA), and was sponsored by the U.S. Department of Education's National Center for Education Statistics (NCES), in cooperation with the U.S. National Science Foundation (NSF) and participating countries. The study, conducted by LessonLab, Inc., was an expansion of the earlier TIMSS Video Study, conducted in 1995 (see textbox on page 3). The goals of the 1999 study were to develop objective observational measures of classroom instruction, describe patterns of teaching practices within each country, and identify similarities and differences in teaching practices across countries.

Data collection for the TIMSS 1999 Video Study began in 1998 and continued through 2000. During that time, 638 videotapes of eighth-grade mathematics lessons were collected from the seven participating countries: Australia, the Czech Republic, Hong Kong SAR, Japan, the Netherlands, Switzerland, and the United States.¹ In each country, mathematics lessons were randomly selected to be representative of eighth-grade mathematics lessons in that country, and teachers were videotaped for one complete, regular class period. In order to capture the range of content and instructional practices that can take place throughout an entire academic year, different classes were videotaped at different times during the school year and lesson content was not controlled for, except that lessons devoted entirely to testing were not filmed. Once a class had been videotaped, the tape was checked for quality and compliance with procedural standards, transcribed, and finally coded by trained coders. Codes were dropped from the study if minimum reliability standards were not met.

Furthermore, in order to determine the representativeness of the final sample, participating teachers were asked a series of questions address-

ing such topics as their academic preparation, years of teaching experience, and their course load, from which their responses were cross-referenced with other known data sets for validation. The result of this exercise indicated that teachers participating in the video study were indeed generally similar to the population of eighth-grade mathematics teachers in their respective countries.

Results

The results of the 1999 study are described in terms of three broadly-grouped components of a lesson, which were each identified as factors that contribute to students' learning opportunities. The three components are lesson structure, lesson content, and instructional practices. It is important to keep in mind, however, that these three categories are not really independent entities, but instead interact with each other to form an overall composite of mathematics instruction in each of the countries. A sampling of findings from the examination of these various components of mathematics instruction are highlighted below.

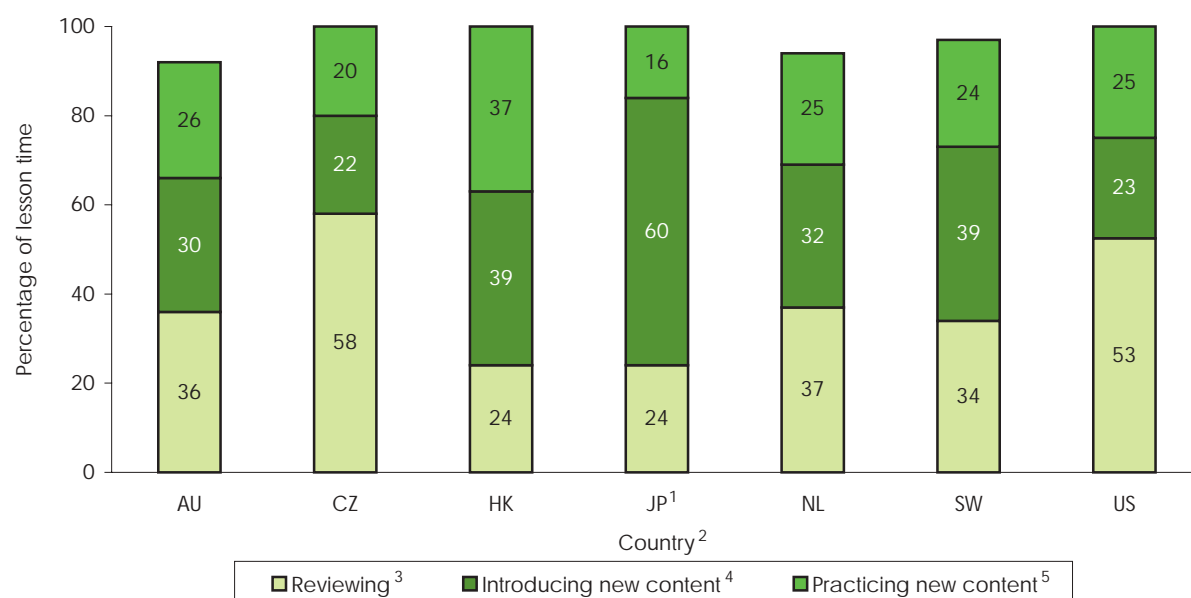
Lesson structure

The organizational structure of students' learning environments may not have a direct influence on learning the way that content and instructional practices may, but it can have an important indirect influence. The structure of a given lesson likely serves to "set the stage" upon which learning will occur, and thus different lesson structures may encourage certain types of learning and hinder others.

There were many similarities across countries with regard to broad features of lesson structure. First, in almost all mathematics lessons in the seven countries, mathematics was taught primarily through problem solving. In fact, in each of the countries, at least 80 percent of lesson time, on average, was spent solving math problems. Second, across all countries, teachers tended to devote some portions of classroom

¹ Hong Kong is a Special Administrative Region (SAR) of the People's Republic of China. For convenience, Hong Kong SAR is referred to as a country. No mathematics classes in Japan were filmed for the 1999 study. Instead, the lessons videotaped for the 1995 TIMSS mathematics study were re-examined and recoded with the expanded 1999 codes.

Figure 1. Average percentage of eighth-grade mathematics lesson time devoted to various purposes, by country: 1999



¹ Japanese mathematics data were collected in 1995.

² AU=Australia; CZ=Czech Republic; HK=Hong Kong SAR; JP=Japan; NL=Netherlands; SW=Switzerland; and US=United States.

³ Reviewing: CZ>AU, HK, JP, NL, SW; US>HK, JP.

⁴ Introducing new content: HK, SW>CZ, US; JP>AU, CZ, HK, NL, SW, US.

⁵ Practicing new content: HK>CZ, JP, SW.

NOTE: For each country, average percentage was calculated as the sum of the percentage within each lesson, divided by the number of lessons. Percentages may not sum to 100 because of rounding and the possibility of coding portions of lessons as "not able

SOURCE: U.S. Department of Education, National Center for Education Statistics, Third International Mathematics and Science Study (TIMSS), Video Study, 1999.

time to reviewing old content, some to introducing new content, and some to practicing new content. Finally, in all countries, classrooms were primarily characterized by two distinct interactive structures: whole-class discussion and individual student work.

Despite these basic similarities, however, there were some interesting variations in lesson structure in different countries. For example, although teachers in all countries devoted some class time to review and some to new content, teachers in the Czech Republic (and to a lesser extent, the United States) tended to emphasize reviewing previously learned content, whereas teachers in Japan emphasized introducing new

content more than in other countries and teachers in Hong SAR emphasized practicing new content more than in three of the other countries (see figure 1).

Pedagogical features related to lesson clarity and flow also varied across countries. In the Czech Republic, for example, teachers were more likely than teachers in all other countries except Japan to explicitly state for their students the specific goals of the lesson. Lessons in the Czech Republic also tended to fall relatively low on measures of the occurrence of potential interruptions to lesson flow (such as outside interruptions and non-mathematical insertions in the lesson). The Netherlands, by contrast, showed the opposite

profile in terms of both goal statements and outside interruptions.

When various elements of lesson structure are examined as a whole, Japan and the Netherlands showed comparatively distinct patterns. Japanese lessons focused on presenting new content, working together as a whole class on only a few problems, and spending a considerable length of time to solve each problem. In the Netherlands, by contrast, individual work was much more common, with eighth-grade students spending a larger percentage of time working on review problems or newly assigned problems.

Lesson content

Lesson content can be broken down into many elements at different levels of detail, from the general topics covered to the specific nature or level of complexity of the actual problems included in the lessons. In the Video Study, researchers focused on the nature and complexity of problems and their relation to the lesson rather than on specific curricular content. Analyses showed that six of the countries were statistically similar with respect to these features, while the seventh country (Japan) showed some dissimilarity.

Close to 15,000 mathematics problems were captured on the videotapes. These problems were subsequently coded and ascribed one of three levels of procedural complexity: low, medium, and high.² Among all of the countries excluding Japan, no differences were detected in the percentages of problems that were of low or high complexity. Mathematics problems covered by teachers in Japanese lessons, however, were of higher complexity relative to those covered in other countries. Thirty-nine percent of problems per lesson, on average, were of high complexity in Japan, while, in the other six countries, 12 percent or fewer problems per lesson,

on average, were of high procedural complexity. The converse was also true: in Japan, 17 percent of the problems were of low complexity, compared to upwards of 63 percent in the other countries (see figure 2).

In addition, the math problems covered in Japanese lessons tended to use, extend, or elaborate on the solution of a previous problem more often than those covered in the lessons in the six other countries. Japanese teachers also presented a lower percentage of repetition problems (problems that were conceptually similar to the previous problem covered in the lesson) compared to the other countries.

The somewhat unique features of Japanese lesson content seem to correspond with the Japanese lesson structure emphasizing introducing new content and devoting large amounts of time to relatively few math problems.

Instructional practices

The Video Study also found that the ways in which teachers and students worked on mathematics content during the lesson varied significantly across countries. In particular, the instructional practices of Dutch and Japanese teachers differed from those of teachers in many of the other countries.

For example, mathematics lessons in the Netherlands tended to emphasize real-world applications of mathematics problems more frequently than lessons in most other countries. Forty-two percent of mathematics problems presented in the Netherlands were connected to real-world situations, compared to between 9 and 27 percent in other countries. Furthermore, when students were assigned a set of problems to be worked on individually in the Netherlands, they were much less likely than their peers in other countries to be subsequently engaged in a whole-class discussion or presentation of the solutions to those problems. In fact, such discussions occurred in 16 percent of possible instances in Dutch classrooms, compared to 76 percent in the Czech Republic.

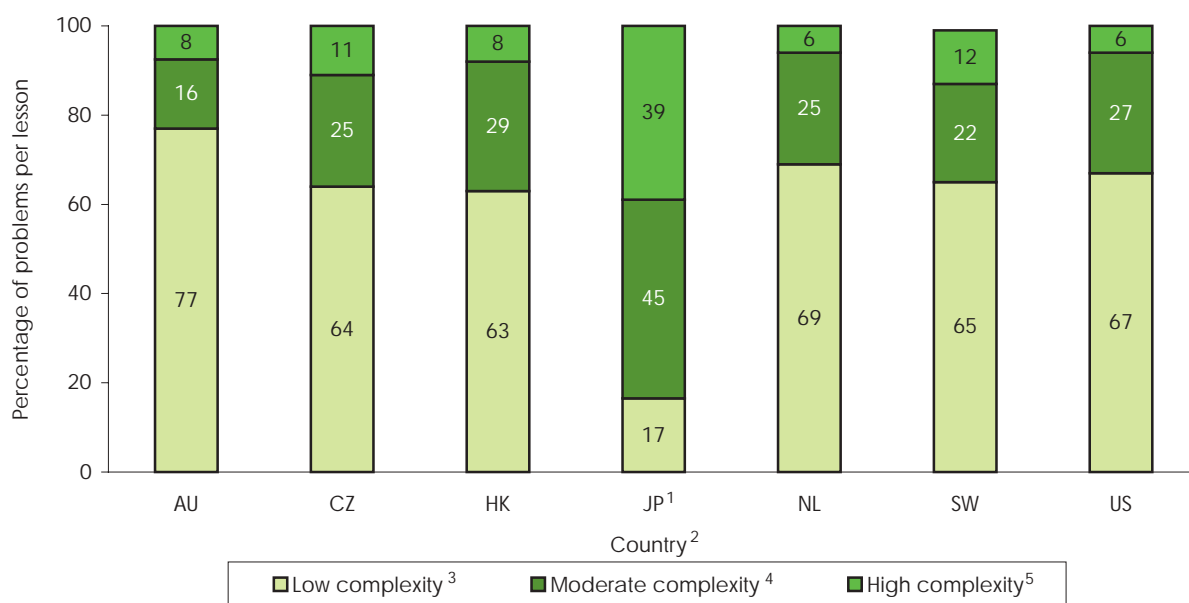
² In general, low complexity problems were those that required four or fewer distinct decisions by a student to solve, while high complexity problems required more than four decisions as well as the completion of at least two sub-problems to solve.

In Japan, teachers presented problems that fell into the category of “applications” (i.e., problems that ask students to apply procedures they have learned in one context to solve problems presented in a different context) more often than teachers in any other country except for Switzerland. Presentations and examinations by students of alternative solution methods for solving a given problem were also more frequent in Japan than in some of the other countries, although overall they occurred relatively infrequently in all countries. Furthermore, in Japan, a higher percentage of problems were summarized by a teacher and stated in a way that emphasized constructing relationships between mathematical facts, ideas, or procedures than

they were in most other countries. By contrast, teachers in Australia and the United States were less likely than teachers in the other countries to summarize problems in this way.

In terms of certain other features of instructional practices, however, countries tended to show relatively similar patterns. For instance, a textbook or worksheet was incorporated into nearly all of the mathematics lessons in each country. As another example, across all countries, teachers spoke more words per lesson relative to their students. In fact, there were very few significant differences across countries in the number of words spoken by teachers for every one word spoken by a student; the exception to this is

Figure 2. Average percentage of eighth-grade mathematics problems per lesson at each level of procedural complexity, by country: 1999



¹ Japanese mathematics data were collected in 1995.

² AU=Australia; CZ=Czech Republic; HK=Hong Kong SAR; JP=Japan; NL=Netherlands; SW=Switzerland; and US=United States.

³ Low complexity: AU, CZ, HK, NL, SW, US > JP.

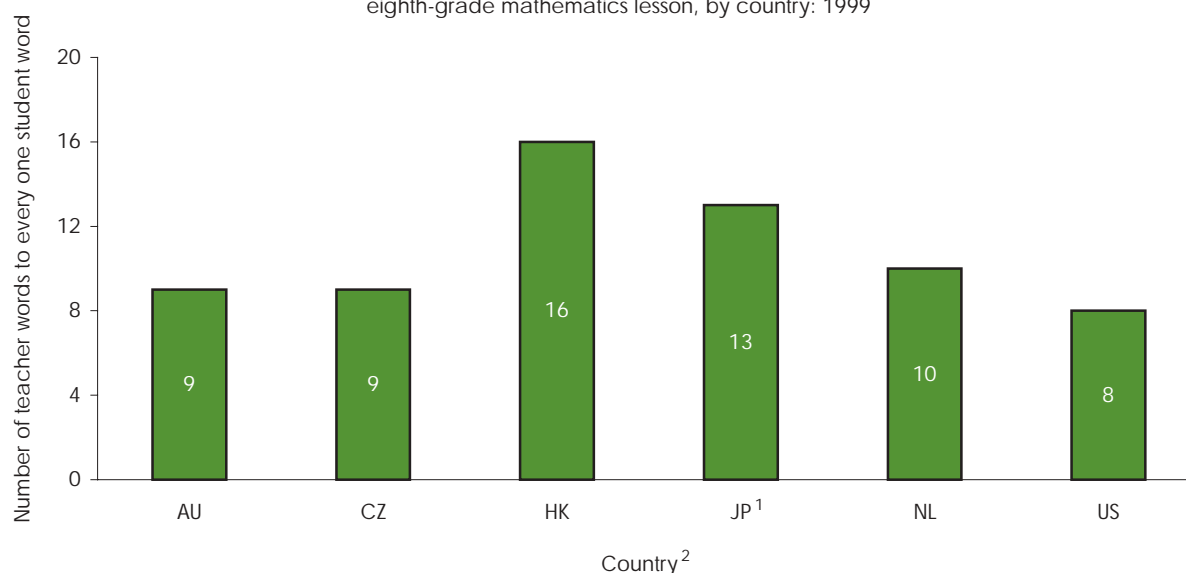
⁴ Moderate complexity: HK > AU; JP > AU, SW.

⁵ High complexity: JP > AU, CZ, HK, NL, SW, US.

NOTE: Percentages may not sum to 100 because of rounding. For each country, average percentage was calculated as the sum of the percentage within each lesson, divided by the number of lessons.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Third International Mathematics and Science Study (TIMSS), Video Study, 1999.

Figure 3. Average number of teacher words to every one student word per eighth-grade mathematics lesson, by country: 1999



¹ Japanese mathematics data were collected in 1995.

² AU=Australia; CZ=Czech Republic; HK=Hong Kong SAR; JP=Japan; NL=Netherlands; and US=United State.

NOTE: HK>AU, CZ, US. Analyses based on English transcripts. English transcripts of Swiss lessons were not available for text analyses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Third International Mathematics and Science Study (TIMSS), Video Study, 1999.

Hong Kong (16:1), where the ratio of teacher-spoken words for every student-spoken word was higher than in Australia (9:1), the Czech Republic (9:1) and the United States (8:1) (see figure 3).

Summary of findings

Overall, the countries participating in the TIMSS 1999 Video Study shared some general features of eighth-grade mathematics teaching, although there were many differences as well. Japan varied from all other countries most often, on 15 percent of the measures analyzed in the report, followed by the Netherlands, which differed from all other countries on 9 percent of the measures. As the results indicated, the differences in Japanese lessons from lessons in other countries were generally based on types of problems included in the lesson (i.e., content) and instructional practices. In the Netherlands, the differences mostly

had to do with lesson structure and instructional practices.

Although they may not have been as marked, each of the other countries exhibited differences from some of the other countries on certain features of mathematics lessons as well. These findings indicate that, among this group of relatively high-achieving countries in mathematics, there is no one common method of teaching mathematics, and instead there are many variations in the way that mathematics is presented.

A full analysis of the results from the mathematics portion of the study can be found in the recently released report *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study* (NCES 2003-013), available on the Web at <http://nces.ed.gov/pubs2003/2003013.pdf>. Additional information and readings can be found at <http://www.lessonlab.com>.

Network A

Network A last met on March 19-21 in Washington, D.C. At the meeting, members discussed Network A's long-term data strategy and the final revisions to the *Education at a Glance 2003* indicators, as well as the progress of the ICT expert panel and the Network A/C Task Force on Teaching and Learning.

With regard to the long-term data strategy, the discussion focused both on issues that relate broadly to a data strategy for INES and those that relate more specifically to the future of PISA. Members discussed such questions as: what data do we need to be measuring in the future, what skills will students need, and what does a quality education system look like? The Network A and OECD Secretariats will work together to draft a paper summarizing the discussion thus far and focusing on issues that will continue to be discussed at future meetings of Network A and the BPC.

Members also discussed final revisions to the draft indicators for *EAG 2003*, particularly concerns about accessibility of some of the language and about the multiple measures of engagement in reading used throughout the indicators. In the end, members agreed to submit five indicators for the chapter on the output of education systems and the impacts of learning: mean performance and variation in fourth-grade students' reading literacy; profiles of 15 year-old readers; 15-year-olds' engagement in reading; 15-year-olds' self-regulated learning; and gender differences in student performance. A framework for the Network's future contributions was also discussed, and a specific proposal for indicators for *EAG 2004* will be sent for members' review over the summer.

Regarding other matters, the progress of the ICT expert panel in developing a framework for assessing ICT literacy in PISA 2006 was presented and supported by members. Members discussed

issues related to funding for the proposed feasibility study, and the OECD Secretariat agreed to send a letter to Network A members requesting additional funds for the study. In addition, comments were given on the third draft of the conceptual framework of the Network A-C Task Force on Teaching and Learning (see textbox on page 10). Members generally supported the framework and suggested further development work to elaborate the relationship of teaching and student outcomes.

The next meeting of Network A will be in Lisbon Portugal on October 16-17, followed by a meeting of the BPC on October 20-22.

Network B

Network B last met on February 3-5 in Madrid, Spain. Eighteen countries were represented, as well as the OECD Secretariat and delegates from Eurostat. At this meeting, Network members discussed indicators and development work in five key areas: educational attainment; continuing education and training (CET); equity, including social outcomes; transition from education to work; and rates of return to education.

OECD presented the indicators on educational attainment for *EAG 2003*, and members also discussed data collection for *EAG 2004*. It was decided that countries would confirm national data in the *EAG 2003* tables and charts by February. Furthermore, members affirmed that, in the future, attainment issues should remain the responsibility of Network B. The status of indicators dealing with the transition from school to work was also presented. It was announced that data requests for 2004 would be sent out in February, with an April deadline for responses. With regard to the development of indicators on equity, it was decided that development efforts should continue, with a greater emphasis in the future on the equity of educational outcomes. However, a separate chapter about equity will not be included in *EAG 2003*. The develop-

ment of indicators of social outcomes of human capital development was also addressed. Because of limited availability of existing surveys and examples, additional development work and discussion will be required, which members supported. Regarding the topic of rate of return to education, Network members discussed issues related to data collection and presentation, and decided to return to this topic at a later date.

With regard to other matters, members reviewed a report on the development of a CET module, which is to serve as a set of guidelines for the development of internationally comparable indicators on CET based on data collected via household surveys. Countries were invited to submit comments on the open issues in the report. Network B also dis-

Network A - C Task Force on Teaching and Learning

Early in 2002, Networks A and C launched a Task Force to develop a long-term data strategy to increase the information available to the INES project on teachers, teaching, and the impact that teachers can have on student learning. This Task Force on Teaching and Learning grew out of the Network C subgroup on teacher indicators and the sustained interest of members of both Networks A and C to learn more about the comparative differences regarding teacher training and development, conditions of employment, teacher's roles and expectations, teaching styles and practices, teacher quality, and the interaction of teaching and learning.

The Task Force, to which 11 countries send expert representatives, has now met three times, with the most recent meeting in May 2003 in Copenhagen. The group has a working conceptual framework, which places the interaction of teacher and student - teaching and learning - at the center of a multilevel framework, analogous to the INES levels of system, school, class, and individual. The framework further identifies 7 categories of inquiry across these levels, including: teaching and learning activities, teacher characteristics and antecedents, student characteristics, ecology of the classroom, teaching and learning relevant school policies and antecedents, teacher workforce characteristics, and policies and antecedents to maintain a high-quality teaching force. The next step for the Task Force is to develop a strategy for how to prioritize and collect information related to the conceptual framework within the INES context. One short-term activity that is being discussed is a teacher survey that would examine the "teacher as worker," seeking information on working conditions, impediments to teaching, and attitudes and expectations. A longer-term activity that also is being discussed is to conduct further development work to support an innovative study that would elaborate the connections between teaching and learning. This strategy paper, covering short- and long-term perspectives, is currently being drafted and will be presented to the Networks at their respective meetings in October and November, after which the Task Force will meet again.

cussed the development work on Young Adults with Low Levels of Education (YALLE), a data collection effort coordinated by the Swedish Secretariat that surveyed individuals age 20 to 24 years-old who have not attained ISCED level 3 and who are enrolled neither in an education nor a work-study program. Based on results from the pilot study in 8 countries, members had decided at their previous meeting to continue with YALLE and recommended that it should be incorporated in regular OECD data collections. At this meeting, it was decided that Sweden would conduct a new

data collection using the pilot data collection as a template. Finally, stemming from a presentation on PISA-L, members decided that a small

task group would work on suggestions for harmonizing school leaver surveys and longitudinal surveys.

Network C

Network C last met in Copenhagen on May 14–16. Network C’s main activities have focused on reviewing the work of the joint Network A/C Task Force on Teaching and Learning, identifying priorities for teacher indicators, planning for system-level indicators, planning for EAG 2004 generally, overseeing reporting from the International Survey of Upper Secondary Schools (ISUSS), and discussing other areas for development.

Regarding the work of the Task Force (see the foregoing text box), Network C, like Network A, endorsed the draft conceptual framework as a basis for further operational work. Furthermore, they recognized the suggested work to expand system-level indicators and develop a teacher survey as being naturally within the domain of Network C and supported additional work related to these activities.

Members also discussed a variety of issues related to planning for future indicators. They discussed the results of a survey of members on their priorities for developing teacher and teaching-related indicators. Fourteen countries had responded to this survey, and at this meeting, members concluded that the results could be used to inform the expansion of system-level indicators as discussed by the Task Force. The Network also discussed the locus of decision-making indicators and decided that members would informally pilot test the survey instrument for ISCED levels 1–3 (general and vocational) in order to help the Network C Secretariat finalize the manual and instructions for the main data collection. The survey will remain unchanged from previous versions in order to allow for time series comparisons. This questionnaire, along with PIRLS, PISA 2000 thematic reports, and ISUSS will be the main sources for potential indicators for EAG 2004. The Network also will

develop its three core indicators (teachers’ working and teaching time, instructional time, and salaries).

Related to ISUSS, which collected data in late 2001/early 2002, members reviewed the OECD’s outline for the final report and heard a report from the subgroup that was involved in reviewing the report. Members agreed to accept the initial report subject to the inclusion of the reviewing subgroup’s comments. They also agreed with the replication of parts of the report for EAG 2003 indicators. OECD will release the report electronically in mid-July and a print copy in September.

Finally, Network C discussed two other topics of a more experimental nature. Since countries are interested in potentially using case studies to collect information and describe monitoring and evaluation systems, the Network C Secretariat will collect some preliminary information on such systems. Members also heard a presentation about educational culture and its potential importance in school performance, but noted the difficulties in developing statistical measures related to this topic.

The next meeting of Network C will be held in November in Korea.

BPC Update

The PISA Board of Participating Countries (BPC) last met in Mexico City in March 2003. The meeting agenda was quite full, with items related to PISA 2000, as well as to the current and future cycles. Related to PISA 2000, members discussed the status of various thematic reports, including on student approaches to learning, student disaffection, and social background. With regard to PISA 2003, the BPC discussed dissemination of results—both in reporting on national experiences during PISA 2000 and in reviewing a proposal for a dissemination strategy for the 2003 results at the international level. The latter also included a review of the outline for the international report, on which there was

a rich discussion. Related to PISA 2006 and beyond, members made final comments on the terms of reference for the next cycle; provided comments on the progress of the expert group in developing the science literacy framework, and adopted criteria and procedures for the participation of non-OECD countries in future PISA

cycles. Finally, the BPC had a preliminary discussion on the long-term development of PISA, which resulted in support for drafting a strategy paper, in conjunction with Network A, for further discussion at future meetings. The next BPC meeting will be held in Lisbon on October 20-22.

Country Highlight: United States

This article was prepared in conjunction with Marilyn Binkley of the U.S. National Center for Education Statistics' Assessment Division.

This article presents a brief overview of the education system in the United States and describes activities to assess student performance at the state and national levels, taking into consideration changes that are occurring as a result of the education legislation passed last year.

A key feature of the U.S. system is that there is no national school system; the United States has a decentralized system of education. The ultimate authority to create and administer educational programs, as codified in the Constitution, rests with the 50 states, and most states have further delegated the authority to operate schools to local governments and educational districts. There are no national laws addressing a prescribed curriculum, the establishment or recognition of institutions, the recognition of degrees or professions, the governance of institutions, or the legal status of students or faculty.

However, although its role is limited, the federal government maintains authority through the U.S. Department of Education for several important activities, including:

- ensuring equity;
- supporting state and local educational improvement efforts;



- providing financial support for postsecondary education;
- helping to make education a national priority; and
- conducting research and gathering statistics, including information to assess how well the overall education system is performing.

We will return to this last responsibility after a brief description of education in the United States.

Brief Overview of Elementary and Secondary Education in the U.S.

Elementary and secondary education in the United States generally spans 12 academic years, or grades. The academic year usually lasts 180 days, with classes in session September to June. However, some educational districts in the United States are now experimenting with year-round schooling, as well. For all but the youngest children, the school day lasts about 7 hours, from approximately 8:00 a.m. to 3:00 p.m. All educational districts require school attendance, although the ages for compulsory schooling may vary slightly by district. Generally, school attendance is required between the ages of 7 and 16.

Thirty-four million children are enrolled in public elementary and lower secondary schools (kindergarten through eighth grade) and 13.5 mil-

lion students in public upper secondary schools (grades 9 through 12). This represents about 89 percent of total enrollment. The other 11 percent of children, 6 million of them, are enrolled in private schools. Private schools may be religiously affiliated and receive most of their funding from student tuition and private donations. Public schools receive about 7 percent of their funding from the federal government, including for compensatory programs aimed at low-income students and schools, and the majority from state and local governments, which provide about 50 and 43 percent, respectively. There are 93,000 public schools in 15,000 school districts in the United States. The number of private schools across the country is 27,000.

'No Child Left Behind' and Assessment in the U.S.

In January 2002, the U.S. Congress passed and President Bush signed the *No Child Left Behind* (NCLB) Act. This new law reauthorized, with some important changes, federal support for elementary and secondary education in the United States. NCLB is built on four basic principles of education reform: stronger accountability for results, increased flexibility and local control, expanded options for parents with children in failing schools, and an emphasis on proven teaching methods. Of greatest interest for this article is the first principle, stronger accountability, which is impacting state testing activities, as well as the national assessment program.

Student Testing in the States

For at least the past decade, U.S. states have been engaged, individually, in efforts to develop state-wide standards for learning. These endeavors have been accompanied by efforts to implement testing programs¹—in the best cases, the state tests are aligned to the standards; in other cases,

the tests are partially aligned to standards or culled from commercially available tests (i.e., “off-the-shelf”).

Currently, all 50 states have some kind of state-wide testing program in one or more subjects such as reading, writing, language arts, mathematics, science, and social studies. Most of these state tests are administered annually, and fourth-, eighth-, and eleventh-graders are the most commonly tested populations—serving in a very loose way as indicators of the cumulative achievement at the end of primary school, middle (lower secondary) school, and high (upper secondary) school. As suggested previously, the degree of alignment of the tests with standards varies from state-to-state, as do the uses to which the resulting information is put.

With NCLB, however, the development of standards and aligned testing now are a requirement for all states in exchange for their federal education funds. While NCLB generally provides states with greater flexibility in administering those funds, it is fairly prescriptive with regard to the accountability aspect, requiring that:

- All states must have standards for all grades for reading and mathematics; and all states must, by the 2005-06 school year, develop standards for all grades for science, as well.
- All states must track students' progress towards those standards through testing programs. Beginning in the 2002-03 school year, states were required to test all students in reading and mathematics in at least one grade from each of the following bands of grades: 3-5, 6-9, and 10-12. Then, beginning in the 2005-06 school year, states must test all students in reading and mathematics in *every* grade, 3 to 8. Science must be added by 2007-08.
- Finally, all states must report their results publicly in state (including district-by-dis-

¹ The terminology individual states use to describe such programs varies from state to state. We use the word “testing” (versus, e.g., “assessment”) because, although

results often are used in the aggregate to evaluate system performance, they also are used for tracking the progress of *individual* students. Thus “test” is more accurate.

trict) and district (including school-by-school) “report cards,” as well as provide information broken down by students’ race/ethnicity, socio-economic status, and disability and language proficiency status, in order to monitor, in particular, the progress of students from disadvantaged groups.

NCLB also allows penalties to be assigned for districts and schools that fail to make adequate yearly progress towards their standards, and gives parents the right to transfer their child to a new school if their child’s current school is consistently deemed to be failing.

In summary, NCLB aims to accelerate and broaden (most) states’ efforts to implement testing programs aligned to learning standards and, in some way, to standardize (at least procedurally, if not in content) the period and scope of testing conducted across states.

National Assessment

Since 1969, the United States has had a national assessment program. Also known as the “Nation’s Report Card,” the National Assessment of Educational Progress (NAEP) provides the only nationally representative and continuing assessment of what students in the United States know and can do. NAEP has three main components:

- National NAEP, which provides information on student performance for the nation overall and by major geographic region;
- State NAEP, begun in 1990, which provides information on student performance at the state level; and
- Long-term trends, which was begun in order to obtain information on changes in U.S. students’ achievement over time.²

² The long-term trend component differs from the national and state components because the frameworks and instruments for the former do not evolve based on changes in curricula or in educational practices (i.e., in order to accurately measure differences over time), whereas the frame-

As these components indicate, NAEP has evolved over the years to meet different information needs. With the enactment of NCLB, it will continue to evolve to reflect the federal government’s increasing focus on educational accountability. Some aspects of the overall program design remain the same (e.g., the three-pronged approach to assessment); other aspects change significantly (e.g., periodicity and the subjects considered).

Changes for National NAEP

Up to the present, national NAEP has periodically assessed a nationally representative sample of students in the 4th, 8th, and 12th grades in a host of subjects including civics, U.S. history, geography, the arts, writing, and, most frequently, reading, mathematics, and science.

From 2003 forward, national NAEP will be regularized according to certain key subject areas. National NAEP will assess students in grades 4 and 8 in mathematics and reading every two years; and it will assess students in grade 12 in these two subjects every four years. National NAEP will (funding permitting) continue to cover additional subjects, albeit a slightly different set of them, including science, civics, U.S. history, geography, writing, the arts, and—new—world history, economics, and foreign language. There will be some national assessment (whether of key or other subjects) occurring each calendar year. (See also Table 1.)

Changes for State NAEP

Up to the present, state NAEP has assessed representative samples of students in the states that chose to participate (generally 40-45 states). Mathematics, science, and reading are the subjects that have been assessed with the most frequency, roughly on opposite cycles every four years (e.g., math and science in 1996 and 2000,

works and instruments for national and state NAEP do change (i.e., in order to best reflect current practice or goals). Therefore, it is not possible to compare results from the long-term trend assessment with those from national and state NAEP.

reading in 1994, 1998, 2002, etc.). The samples used in state NAEP have been separate from the sample drawn for national NAEP, which under its current design cannot support disaggregating by state.

The major changes for state NAEP are that: (1) it will now be mandatory for states to participate in the 4th and 8th grade reading and mathematics assessments; (2) the state data will be generated from the same sample as drawn for national NAEP; and, building off this, (3) it is being brought into alignment with the schedule for national NAEP. As such, beginning in 2003, mathematics and reading will be assessed and broken down at the state level every two years for students in grades 4 and 8. Science and writing will be assessed every four years in these grades.

Changes for Long-Term Trends

Up to the present, the long-term trend component of NAEP has assessed mathematics, reading, and science at ages 9, 13, and 17, and additionally has assessed writing at grades 4, 8, and 12. In the recent past, assessments have occurred roughly every 2 years.

In contrast to the changes expected for state and national NAEP, the key changes for long-term trends are aimed at scaling-back this component and increasing overall efficiency (a change that had already been underway prior to NCLB). Given what is observed to be a very gradual pace of change in education outcomes, the period of assessment will now be every 4 years; the science and writing components will be phased out altogether.

Table 1: Preliminary Schedule for NAEP: 2003 – 2012

Year	National	State	Long-Term
2003	Reading, Mathematics	Reading, Mathematics	
2004	Foreign language (12)		Reading, Mathematics
2005	Reading, Mathematics, Science	Reading, Mathematics, Science	
2006	World History, Economics, Civics		
2007	Reading, Mathematics, Writing	Reading, Mathematics, Writing	
2008	Arts		Reading, Mathematics
2009	Reading, Mathematics, Science	Reading, Mathematics, Science	
2010	U.S. history, Geography		
2011	Reading, Mathematics, Writing	Reading, Mathematics, Writing	
2012	Foreign language, Civics		Reading, Mathematics

SOURCE: <http://nces.ed.gov/nationsreportcard/>

This newsletter is published under the auspices of Network A. Network A, which is primarily concerned with indicators of learner outcomes, is one of several working groups that are part of OECD's international Indicators of Education Systems (INES) Project. The newsletter is prepared by Eugene Owen (Network A Chair) and Maria Stephens and Christina Stearns of the American Institutes for Research with contributions from Network A members. The newsletter was designed by Charmaine Llagas of the American Institutes for Research. Comments on the newsletter may be sent to mstephens@air.org.

Network A Representatives

Mr. Eugene Owen (<i>Chair</i>) International Studies Program National Center for Education Statistics UNITED STATES	Ms. Jacqueline Levasseur Ministère de l'éducation nationale FRANCE	Mr. Jules Peschar University of Groningen NETHERLANDS
Ms. Wendy Whitham Department of Education, Science and Training AUSTRALIA	Mr. Jochen Schweitzer Der Senator für Bildung und Wissenschaft Der Freien Hansestadt Bremen GERMANY	Ms. Lynne Whitney Ministry of Education NEW ZEALAND
Mr. Helmut Bachmann Federal Ministry of Education, Science and Culture AUSTRIA	Ms. Evangelia Varnava-Skoura Ministry of Education GREECE	Ms. Anne-Berit Kavli Norwegian Board of Education NORWAY
Mr. Jürgen Horschneegg Federal Ministry of Education, Science and Culture AUSTRIA	Mr. Benedek Péter Tóta Ministry of Education HUNGARY	Ms. Glória Ramalho Gabinete de Avaliação Educacional (GAVE)Ministério da Educação PORTUGAL
Ms. Christiane Blondin Service de Pédagogie expérimentale de l'Université de Liège BELGIUM	Mr. Gerry Shiel Educational Research Centre St. Patrick's College IRELAND	Mr. Vladislav Rosa Director of Central School Inspection SLOVAK REPUBLIC
Mr. Luc Van de Poele Department of Education Universiteit Gent BELGIUM	Ms. Chiara Croce Direzione Generale Scambi Culturali Ministero della Pubblica Istruzione ITALY	Mr. Guillermo Gil Instituto Nacional de Calidad y Evaluación Ministerio de Educación y Ciencia SPAIN
Mr. Michael O'Gorman, Curriculum & Assessment Policy Branch Ministry of Education Ontario CANADA	Mr. Ryo Watanabe Department of International Education National Institute for Educational Research JAPAN	Ms. Anna Barklund Swedish National Agency for Education SWEDEN
Mr. Jerry Mussio Centre for Education Statistics Statistics Canada CANADA	Mr. Myungjoon Lee International Research and Cooperation Korea Institute of Curriculum and Evaluation KOREA	Ms. Anita Wester National Agency for Education SWEDEN
Ms. Pavla Zieleniecova Institute for Information on Education CZECH REPUBLIC	Ms. Iris Blanke Ministère de l'éducation nationale LUXEMBOURG	Mr. Erich Ramseier Amt fuer Bildungsforschung SWITZERLAND
Mr. Jørn Skovsgaard Ministry of Education DENMARK	Mr. Fernando Córdova Secretaria de Educacion Publica MÉXICO	Sevki Karaca Directorate for Research and Development TURKEY
Ms. Pirjo Linnakylä Institute for Educational Research University of Jyväskylä FINLAND	Mr. Paul van Oijen Ministry of Education Science and Research NETHERLANDS	Mr. Jason Tarsh Department for Education and Skills UNITED KINGDOM
		Ms. Mariann Lemke International Studies Program National Center for Education Statistics UNITED STATES